

#3

OIEP

## RAW SEQUENCE LISTING

DATE: 12/26/2001

PATENT APPLICATION: US/09/884,948

TIME: 16:19:42

Input Set : N:\Crif3\RULE60\09884948.raw

Output Set: N:\CRF3\12262001\I884948.raw

1 <110> APPLICANT: YOKOYAMA, KEIICHI  
 2 NAKAMURA, NAMI  
 3 MIWA, TETSUYA  
 4 SEGURO, KATSUYA  
 5 <120> TITLE OF INVENTION: PROCESS FOR PRODUCING MICROBIAL TRANSGLUTAMINASE  
 6 <130> FILE REFERENCE: 0010-0937-0  
 7 <140> CURRENT APPLICATION NUMBER: 09/884,948  
 8 <141> CURRENT FILING DATE: 2001-06-21  
 9 <150> PRIOR APPLICATION NUMBER: 09/448,310  
 10 <151> PRIOR FILING DATE: 1999-11-24  
 11 <160> NUMBER OF SEQ ID NOS: 62  
 12 <170> SOFTWARE: PatentIn Ver. 2.0  
 14 <210> SEQ ID NO: 1  
 15 <211> LENGTH: 331  
 16 <212> TYPE: PRT  
 17 <213> ORGANISM: Artificial Sequence  
 18 <220> FEATURE:  
 19 <223> OTHER INFORMATION: Description of Artificial  
 20 Sequence: TRANSGLUTAMINASE  
 21 <400> SEQUENCE: 1  
 22 Asp Ser Asp Asp Arg Val Thr Pro Pro Ala Glu Pro Leu Asp Arg Met  
 23 1 5 10 15  
 24 Pro Asp Pro Tyr Arg Pro Ser Tyr Gly Arg Ala Glu Thr Val Val Asn  
 25 20 25 30  
 26 Asn Tyr Ile Arg Lys Trp Gln Gln Val Tyr Ser His Arg Asp Gly Arg  
 27 35 40 45  
 28 Lys Gln Gln Met Thr Glu Glu Gln Arg Glu Trp Leu Ser Tyr Gly Cys  
 29 50 55 60  
 30 Val Gly Val Thr Trp Val Asn Ser Gly Gln Tyr Pro Thr Asn Arg Leu  
 31 65 70 75 80  
 32 Ala Phe Ala Ser Phe Asp Glu Asp Arg Phe Lys Asn Glu Leu Lys Asn  
 33 85 90 95  
 34 Gly Arg Pro Arg Ser Gly Glu Thr Arg Ala Glu Phe Glu Gly Arg Val  
 35 100 105 110  
 36 Ala Lys Glu Ser Phe Asp Glu Glu Lys Gly Phe Gln Arg Ala Arg Glu  
 37 115 120 125  
 38 Val Ala Ser Val Met Asn Arg Ala Leu Glu Asn Ala His Asp Glu Ser  
 39 130 135 140  
 40 Ala Tyr Leu Asp Asn Leu Lys Lys Glu Leu Ala Asn Gly Asn Asp Ala  
 41 145 150 155 160  
 42 Leu Arg Asn Glu Asp Ala Arg Ser Pro Phe Tyr Ser Ala Leu Arg Asn  
 43 165 170 175  
 44 Thr Pro Ser Phe Lys Glu Arg Asn Gly Gly Asn His Asp Pro Ser Arg  
 45 180 185 190  
 46 Met Lys Ala Val Ile Tyr Ser Lys His Phe Trp Ser Gly Gln Asp Arg  
 47 195 200 205  
 48 Ser Ser Ser Ala Asp Lys Arg Lys Tyr Gly Asp Pro Asp Ala Phe Arg

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/884,948

DATE: 12/26/2001

TIME: 16:19:42

Input Set : N:\Crif3\RULE60\09884948.raw

Output Set: N:\CRF3\12262001\I884948.raw

```

49          210          215          220
50  Pro Ala Pro Gly Thr Gly Leu Val Asp Met Ser Arg Asp Arg Asn Ile
51  225          230          235          240
52  Pro Arg Ser Pro Thr Ser Pro Gly Glu Gly Phe Val Asn Phe Asp Tyr
53          245          250          255
54  Gly Trp Phe Gly Ala Gln Thr Glu Ala Asp Ala Asp Lys Thr Val Trp
55          260          265          270
56  Thr His Gly Asn His Tyr His Ala Pro Asn Gly Ser Leu Gly Ala Met
57          275          280          285
58  His Val Tyr Glu Ser Lys Phe Arg Asn Trp Ser Glu Gly Tyr Ser Asp
59          290          295          300
60  Phe Asp Arg Gly Ala Tyr Val Ile Thr Phe Ile Pro Lys Ser Trp Asn
61  305          310          315          320
62  Thr Ala Pro Asp Lys Val Lys Gln Gly Trp Pro
63          325          330
65 <210> SEQ ID NO: 2
66 <211> LENGTH: 993
67 <212> TYPE: DNA
68 <213> ORGANISM: Artificial Sequence
69 <220> FEATURE:
70 <223> OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC DNA
71 <221> NAME/KEY: CDS
72 <222> LOCATION: (1)..(993)
73 <223> OTHER INFORMATION: IDENTIFICATION METHOD: S
74 <400> SEQUENCE: 2
75  gat tct gac gat cgt gtt act cca cca gct gaa cca ctg gat cgt atg 48
76  Asp Ser Asp Asp Arg Val Thr Pro Pro Ala Glu Pro Leu Asp Arg Met
77  1          5          10          15
78  cca gat cca tat cgt cca tct tat ggt cgt gct gaa act gtt gtt aat 96
79  Pro Asp Pro Tyr Arg Pro Ser Tyr Gly Arg Ala Glu Thr Val Val Asn
80          20          25          30
81  aat tat att cgt aaa tgg caa caa gtt tat tct cat cgt gat ggt cgt 144
82  Asn Tyr Ile Arg Lys Trp Gln Gln Val Tyr Ser His Arg Asp Gly Arg
83          35          40          45
84  aaa caa caa atg act gaa gaa caa cgt gaa tgg ctg tct tat ggt tgc 192
85  Lys Gln Gln Met Thr Glu Glu Gln Arg Glu Trp Leu Ser Tyr Gly Cys
86          50          55          60
87  gtt ggt gtt act tgg gtt aac tct ggt cag tat ccg act aac cgt ctg 240
88  Val Gly Val Thr Trp Val Asn Ser Gly Gln Tyr Pro Thr Asn Arg Leu
89          65          70          75          80
90  gca ttc gct tcc ttc gat gaa gat cgt ttc aag aac gaa ctg aag aac 288
91  Ala Phe Ala Ser Phe Asp Glu Asp Arg Phe Lys Asn Glu Leu Lys Asn
92          85          90          95
93  ggt cgt ccg cgt tct ggt gaa act cgt gct gaa ttc gaa ggt cgt gtt 336
94  Gly Arg Pro Arg Ser Gly Glu Thr Arg Ala Glu Phe Glu Gly Arg Val
95          100          105          110
96  gct aag gaa tcc ttc gat gaa gag aaa ggc ttc cag cgt gct cgt gaa 384
97  Ala Lys Glu Ser Phe Asp Glu Glu Lys Gly Phe Gln Arg Ala Arg Glu
98          115          120          125

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/884,948

DATE: 12/26/2001

TIME: 16:19:42

Input Set : N:\Crf3\RULE60\09884948.raw

Output Set: N:\CRF3\12262001\I884948.raw

```

99      gtt gct tct gtt atg aac cgt gct cta gag aac gct cat gat gaa tct 432
100     Val Ala Ser Val Met Asn Arg Ala Leu Glu Asn Ala His Asp Glu Ser
101           130                      135                      140
102     gct tac ctg gat aac ctg aag aag gaa ctg gct aac ggt aac gat gct 480
103     Ala Tyr Leu Asp Asn Leu Lys Lys Glu Leu Ala Asn Gly Asn Asp Ala
104           145                      150                      155                      160
105     ctg cgt aac gaa gat gct cgt tct ccg ttc tac tct gct ctg cgt aac 528
106     Leu Arg Asn Glu Asp Ala Arg Ser Pro Phe Tyr Ser Ala Leu Arg Asn
107           165                      170                      175
108     act ccg tcc ttc aaa gaa cgt aac ggt ggt aac cat gat ccg tct cgt 576
109     Thr Pro Ser Phe Lys Glu Arg Asn Gly Gly Asn His Asp Pro Ser Arg
110           180                      185                      190
111     atg aaa gct gtt atc tac tct aaa cat ttc tgg tct ggt cag gat aga 624
112     Met Lys Ala Val Ile Tyr Ser Lys His Phe Trp Ser Gly Gln Asp Arg
113           195                      200                      205
114     tct tct tct gct gat aaa cgt aaa tac ggt gat ccg gat gca ttc cgt 672
115     Ser Ser Ser Ala Asp Lys Arg Lys Tyr Gly Asp Pro Asp Ala Phe Arg
116           210                      215                      220
117     ccg gct ccg ggt act ggt ctg gta gac atg tct cgt gat cgt aac atc 720
118     Pro Ala Pro Gly Thr Gly Leu Val Asp Met Ser Arg Asp Arg Asn Ile
119           225                      230                      235                      240
120     ccg cgt tct ccg act tct ccg ggt gaa ggc ttc gtt aac ttc gat tac 768
121     Pro Arg Ser Pro Thr Ser Pro Gly Glu Gly Phe Val Asn Phe Asp Tyr
122           245                      250                      255
123     ggt tgg ttc ggt gct cag act gaa gct gat gct gat aag act gta tgg 816
124     Gly Trp Phe Gly Ala Gln Thr Glu Ala Asp Ala Asp Lys Thr Val Trp
125           260                      265                      270
126     acc cat ggt aac cat tac cat gct ccg aac ggt tct ctg ggt gct atg 864
127     Thr His Gly Asn His Tyr His Ala Pro Asn Gly Ser Leu Gly Ala Met
128           275                      280                      285
129     cat gta tac gaa tct aaa ttc cgt aac tgg tct gaa ggt tac tct gac 912
130     His Val Tyr Glu Ser Lys Phe Arg Asn Trp Ser Glu Gly Tyr Ser Asp
131           290                      295                      300
132     ttc gat cgt ggt gct tac gtt atc acc ttc att ccg aaa tct tgg aac 960
133     Phe Asp Arg Gly Ala Tyr Val Ile Thr Phe Ile Pro Lys Ser Trp Asn
134           305                      310                      315                      320
135     act gct ccg gac aaa gtt aaa cag ggt tgg ccg 993
136     Thr Ala Pro Asp Lys Val Lys Gln Gly Trp Pro
137           325                      330
139 <210> SEQ ID NO: 3
140 <211> LENGTH: 1519
141 <212> TYPE: DNA
142 <213> ORGANISM: Artificial Sequence
143 <220> FEATURE:
144 <223> OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC DNA
145 <221> NAME/KEY: CDS
146 <222> LOCATION: (87)..(1082)
147 <400> SEQUENCE: 3
148      ttccctggtt gacaattaat catcgaacta gttaactagt acgcaagttc acgtaaaaag 60

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/884,948

DATE: 12/26/2001

TIME: 16:19:42

Input Set : N:\Crf3\RULE60\09884948.raw

Output Set: N:\CRF3\12262001\I884948.raw

```

149      ggtatcgcatt agtaaggagg tttaaa atg gat tct gac gat cgt gtt act cca 113
150                                     Met Asp Ser Asp Asp Arg Val Thr Pro
151                                     1           5
152      cca gct gaa cca ctg gat cgt atg cca gat cca tat cgt cca tct tat 161
153      Pro Ala Glu Pro Leu Asp Arg Met Pro Asp Pro Tyr Arg Pro Ser Tyr
154      10           15           20           25
155      ggt cgt gct gaa act gtt gtt aat aat tat att cgt aaa tgg caa caa 209
156      Gly Arg Ala Glu Thr Val Val Asn Asn Tyr Ile Arg Lys Trp Gln Gln
157      30           35           40
158      gtt tat tct cat cgt gat ggt cgt aaa caa caa atg act gaa gaa caa 257
159      Val Tyr Ser His Arg Asp Gly Arg Lys Gln Gln Met Thr Glu Glu Gln
160      45           50           55
161      cgt gaa tgg ctg tct tat ggt tgc gtt ggt gtt act tgg gtt aac tct 305
162      Arg Glu Trp Leu Ser Tyr Gly Cys Val Gly Val Thr Trp Val Asn Ser
163      60           65           70
164      ggt cag tat ccg act aac cgt ctg gca ttc gct tcc ttc gat gaa gat 353
165      Gly Gln Tyr Pro Thr Asn Arg Leu Ala Phe Ala Ser Phe Asp Glu Asp
166      75           80           85
167      cgt ttc aag aac gaa ctg aag aac ggt cgt ccg cgt tct ggt gaa act 401
168      Arg Phe Lys Asn Glu Leu Lys Asn Gly Arg Pro Arg Ser Gly Glu Thr
169      90           95           100           105
170      cgt gct gaa ttc gaa ggt cgt gtt gct aag gaa tcc ttc gat gaa gag 449
171      Arg Ala Glu Phe Glu Gly Arg Val Ala Lys Glu Ser Phe Asp Glu Glu
172      110           115           120
173      aaa ggc ttc cag cgt gct cgt gaa gtt gct tct gtt atg aac cgt gct 497
174      Lys Gly Phe Gln Arg Ala Arg Glu Val Ala Ser Val Met Asn Arg Ala
175      125           130           135
176      cta gag aac gct cat gat gaa tct gct tac ctg gat aac ctg aag aag 545
177      Leu Glu Asn Ala His Asp Glu Ser Ala Tyr Leu Asp Asn Leu Lys Lys
178      140           145           150
179      gaa ctg gct aac ggt aac gat gct ctg cgt aac gaa gat gct cgt tct 593
180      Glu Leu Ala Asn Gly Asn Asp Ala Leu Arg Asn Glu Asp Ala Arg Ser
181      155           160           165
182      ccg ttc tac tct gct ctg cgt aac act ccg tcc ttc aaa gaa cgt aac 641
183      Pro Phe Tyr Ser Ala Leu Arg Asn Thr Pro Ser Phe Lys Glu Arg Asn
184      170           175           180           185
185      ggt ggt aac cat gat ccg tct cgt atg aaa gct gtt atc tac tct aaa 689
186      Gly Gly Asn His Asp Pro Ser Arg Met Lys Ala Val Ile Tyr Ser Lys
187      190           195           200
188      cat ttc tgg tct ggt cag gat aga tct tct tct gct gat aaa cgt aaa 737
189      His Phe Trp Ser Gly Gln Asp Arg Ser Ser Ser Ala Asp Lys Arg Lys
190      205           210           215
191      tac ggt gat ccg gat gca ttc cgt ccg gct ccg ggt act ggt ctg gta 785
192      Tyr Gly Asp Pro Asp Ala Phe Arg Pro Ala Pro Gly Thr Gly Leu Val
193      220           225           230
194      gac atg tct cgt gat cgt aac atc ccg cgt tct ccg act tct ccg ggt 833
195      Asp Met Ser Arg Asp Arg Asn Ile Pro Arg Ser Pro Thr Ser Pro Gly
196      235           240           245
197      gaa ggc ttc gtt aac ttc gat tac ggt tgg ttc ggt gct cag act gaa 881

```

## RAW SEQUENCE LISTING

DATE: 12/26/2001

PATENT APPLICATION: US/09/884,948

TIME: 16:19:42

Input Set : N:\Crf3\RULE60\09884948.raw

Output Set: N:\CRF3\12262001\I884948.raw

```

198      Glu Gly Phe Val Asn Phe Asp Tyr Gly Trp Phe Gly Ala Gln Thr Glu
199      250                      255                      260                      265
200      gct gat gct gat aag act gta tgg acc cat ggt aac cat tac cat gct      929
201      Ala Asp Ala Asp Lys Thr Val Trp Thr His Gly Asn His Tyr His Ala
202                      270                      275                      280
203      ccg aac ggt tct ctg ggt gct atg cat gta tac gaa tct aaa ttc cgt      977
204      Pro Asn Gly Ser Leu Gly Ala Met His Val Tyr Glu Ser Lys Phe Arg
205                      285                      290                      295
206      aac tgg tct gaa ggt tac tct gac ttc gat cgt ggt gct tac gtt atc      1025
207      Asn Trp Ser Glu Gly Tyr Ser Asp Phe Asp Arg Gly Ala Tyr Val Ile
208                      300                      305                      310
209      acc ttc att ccg aaa tct tgg aac act gct ccg gac aaa gtt aaa cag      1073
210      Thr Phe Ile Pro Lys Ser Trp Asn Thr Ala Pro Asp Lys Val Lys Gln
211                      315                      320                      325
212      ggt tgg ccg taatgaaagc ttggatctct aattactgga cttcacacag      1122
213      Gly Trp Pro
214      330
215      actaaaatag acatatctta tattatgtga ttttgtgaca tttcctagat gtgaggtgga 1182
216      gctgatgtat aaggtagatg atgatoctct acgccggacg catcgtggcc gccatcaccg 1242
217      gcgccacagg tgcggttgct ggcgcctata tcgccgacat caccgatggg gaagatcggg 1302
218      ctgcgcaactt cgggctcatg agcgcttggt tcggcgtggg tatggtggca ggccccgtgg 1362
219      ccgggggact gttggggcgcc atctccttgc atgcaccatt ccttgcgggc gcggtgctca 1422
220      acggcctcaa cctactactg ggctgcttcc taatgcagga gtcgcataag ggagagcgtc 1482
221      gagagcccg ctaatgagcg ggcttttttt tcagctg      1519
223 <210> SEQ ID NO: 4
224 <211> LENGTH: 39
225 <212> TYPE: DNA
226 <213> ORGANISM: Artificial Sequence
227 <220> FEATURE:
228 <223> OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC DNA
229 <400> SEQUENCE: 4
230      aattcatcga ttagtaagga ggtttaaaat ggattctga      39
232 <210> SEQ ID NO: 5
233 <211> LENGTH: 41
234 <212> TYPE: DNA
235 <213> ORGANISM: Artificial Sequence
236 <220> FEATURE:
237 <223> OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC DNA
238 <400> SEQUENCE: 5
239      cgatcgtcag aatccatttt aaacctcctt actaatcgat g      41
241 <210> SEQ ID NO: 6
242 <211> LENGTH: 41
243 <212> TYPE: DNA
244 <213> ORGANISM: Artificial Sequence
245 <220> FEATURE:
246 <223> OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC DNA
247 <400> SEQUENCE: 6
248      cgatcgtggt actccaccag ctgaaccact ggatcgtatg c      41
250 <210> SEQ ID NO: 7

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/884,948

DATE: 12/26/2001

TIME: 16:19:43

Input Set : N:\Crf3\RULE60\09884948.raw

Output Set: N:\CRF3\12262001\I884948.raw